Jacqueline Wong, Kevin Nguyen

CS 6301.502

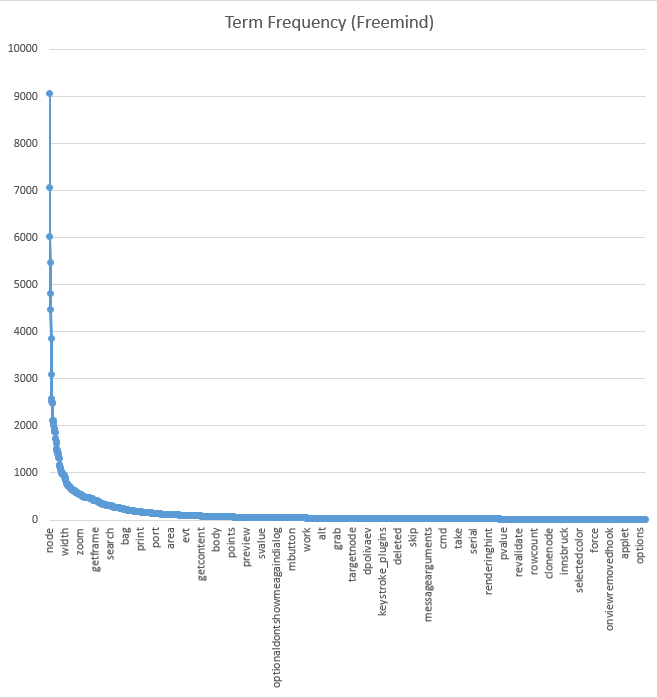
SW Analysis and Comprehension

Team 4

Assignment #1 - Deliverable

**FreeMind**

1. Plot of the distribution of all the terms in the software lexicon. You can use Excel or anything like it. (Excel only picked a few random terms to show since it couldn’t fit all.)



1. List of the 10 most frequent terms in the corpus (report their tf)

node 9062

map 7068

string 6024

action 5481

set 4814

mind 4471

controller 3859

null 3083

free 2588

add 2516

1. List of the 10 terms with the highest document frequency (report their df).

free 488

write 486

copyright 486

version 485

software 485

modify 485

purpose 484

license 484

later 484

term 483

1. List of the 3 documents with the highest number of unique terms. (Highest to Lowest)

MindMapController.java

FreeMindMapController.java

FreeMind.java

1. List of the 3 documents with the lowest number of unique terms.(Lowest to Highest)

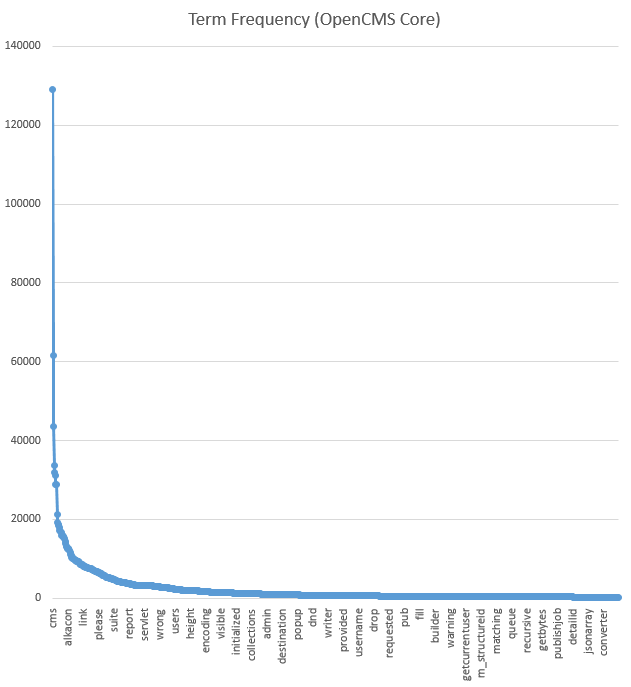
IconInformation.java

SignedScriptTests.java

LoggingOutputStream.java

**OpenCMS Core**

1. Plot of the distribution of all the terms in the software lexicon. You can use Excel or anything like it.



1. List of the 10 most frequent terms in the corpus (report their tf)

cms 129069

string 61694

resource 43468

param 33701

name 31790

opencms 31133

set 28922

list 28816

value 21209

open 19119

1. List of the 10 terms with the highest document frequency (report their df).

org 3128

version 3126

http 3126

copy 3126

www 3125

software 3124

license 3124

imply 3124

distribute 3124

copyright 3124

1. List of the 3 documents with the highest number of unique terms. (Highest to Lowest)

CmsDriverManager.java

CmsDefaultXmlContentHandler.java

CmsSecurityManager.java

1. List of the 3 documents with the lowest number of unique terms.(Lowest to Highest)

CmsRelationValidatorInfoEntry.java

SynchronizedRpcRequest.java

I\_CmsUserAgentInfo.java